January 1979

# Accuracy of First Lactation Versus All Lactation Sire Evaluations by Best Linear Unbiased Prediction

G. R. Ufford
*Cornell University*

C. R. Henderson
*Cornell University*

J. F. Keown
*Cornell University*, jkeown1@unl.edu

L. Dale Van Vleck
*University of Nebraska-Lincoln*, dvan-vleck1@unl.edu

# Accuracy of First Lactation Versus All Lactation Sire Evaluations by Best Linear Unbiased Prediction

G. R. UFFORD[1], C. R. HENDERSON, J. F. KEOWN[2], and L. D. VAN VLECK
Department of Animal Science
Cornell University
Ithaca, NY 14853

## ABSTRACT

Variances of errors of prediction for sire evaluations which included only first records and for those with records of all lactations were compared for bulls of Ayrshire, Guernsey, Jersey, and Brown Swiss breeds used by artificial insemination with daughters having Dairy Herd Improvement records processed at the New York Dairy Records Processing Laboratory. The model for best linear unbiased prediction included fixed effects of sire group and herd-year-season of freshening and random effects of sires within group, sire-by-herd interaction (to account for environmental correlation among paternal sisters), cow within sire and herd, and residual. Variances of solutions for group effects were generally small relative to variances of prediction errors for sire effects. Using all lactation records, however, reduced variances of group solutions by 7 to 14% for groups of sires used artificially and by 20 to 24% for groups used in natural service. Use of all lactation records decreased the variance of prediction error of the sire solutions so that 15 daughters per sire with all lactations gave accuracy equivalent to 25 daughters using only first records; use of all lactation records with 25 daughters gave accuracy equivalent to 40 daughters with only first records. Genetic progress per year from selection of bulls to sire daughters would be expected to be 10 to 15% greater with use of all lactation records than with use of only first lactation records. The comparable increase from selection of bulls to sire replacement bulls would be 3 to 10%. These theoretical increases must be weighed against possible biases from use of records other than first lactation.

## INTRODUCTION

The value of incorporating more than first lactation records is an important consideration in implementation of sire evaluation by Best Linear Unbiased Prediction (BLUP) procedures. Computing costs are increased considerably by incorporation of later lactation records. Acceptance of a procedure for sire evaluation by scientists, educators, and producers is influenced by the use of or omission of later lactation records. Perhaps most importantly, the accuracy of evaluations may be increased or decreased by addition of later lactation records. Considering later lactation records, environmental correlation, or natural service (NS) sires adds considerably to the complexity of computing evaluations as compared to use of first lactation records on daughters sired artificially (AI).

Henderson (4) stated that whenever feasible the way to compare alternative procedures of sire evaluation is by some analytical method. Henderson (1, 2) discussed several criteria appropriate for genetic evaluation including maximizing the probability of correct pairwise ranking and maximizing the mean of a selected group. However, with present statistical knowledge, a workable and yet meaningful criterion is minimum mean squared error, with expectation over all random effects in the model, including the genetic effects to be predicted as well as residual effects.

Mean squared error is the sum of bias squared and error variance of prediction. In problems of mixed models with unknown means it is difficult, if not impossible, to minimize the mean squared error. An alternative is to restrict consideration to unbiased proce-

Received for publication September 25, 1978.
[1] 34516 Tennessee Rd., Lebanon, OR 97355.
[2] Eastern Artificial Insemination Coop., Inc., Ithaca, NY 14850.

dures where minimum squared error is equivalent to minimum prediction error variance. Under assumptions of the specified model, BLUP procedures do minimize prediction error variance. If the assumptions are violated, error variances may be increased or bias may be introduced (3, 4).

The merit of incorporating all lactation records can be considered in terms of error variance of prediction and bias. Since first lactation records are a subset of all lactation records, error variances of prediction will be smaller when all records are used. However, bias may result from inclusion of later records. Incorrect age-season factors could cause bias if all lactation records are used, because sires differ in the proportion of first lactation records. Culling of cows based on their records also will cause bias if variance components are incorrect (3).

Thus, mean squared error may increase or decrease depending on the amount that bias is increased and error variance is decreased when the additional records are included. The primary purpose of this study was to compare error variances of prediction that may be encountered in sire evaluations from using first as compared to all lactation records. These error variances are a function of the design, e.g., distribution of sires across herd-year-seasons, but are not influenced by the actual production records. Thus, field data need be used only to determine representative designs for comparison of error variances of prediction for first vs. all lactation procedures under the assumption that first and later records are measures of the same trait.

## METHODS

The model with sire-by-herd interaction was used as in (8, 9). Briefly:

$$y_{fijklm} = g_i + s_{fij} + h_{kl} + sh_{fijk} + c_{fijkm} + e_{fijklm} \quad [1]$$

where:

- y is an age-season adjusted record,
- g is a fixed effect of sire group,
- s is a random effect of sire within group with $\sigma_e^2/\sigma_s^2 = 8.33$ ($f$ indicates an AI or NS sire),
- h is a fixed effect of herd-year-season ($l$th year-season within the $k$th herd),
- sh is a random effect of sire-by-herd interaction (accounting for the environmental correlation among paternal half-sisters) with $\sigma_e^2/\sigma_{sh}^2 = 3.57$,
- c is a random effect of cow within sire and herd with $\sigma_e^2/\sigma_c^2 = 1.67$, and
- e is a random residual effect with variance $\sigma_e^2$.

The variance ratios 8.33, 3.57, and 1.67 were suggested by Norman (6) because they correspond to $h^2$ of .24, $c^2$ of .14, and intracow repeatability of .50.

Data from the New York Dairy Records Processing Laboratory (DRPL) were used. Up to 10 lactations were accepted on cows that had a first lactation record in the same herd. The data set included records normally used for AI sire evaluation plus records on daughters of NS sires. The AI and NS sires were assigned to separate groups as by DRPL. Records of Ayrshires, Guernseys, Jerseys, and Brown Swiss were studied. The data set for Holsteins was too large to make the calculations needed for the comparisons.

The computing procedures for best linear unbiased prediction for the model are described in (8). The sire evaluation is defined as $\hat{g}_i + \hat{s}_{ij}$.

After absorption of equations for cows, sire-by-herd, herd-year-season, and NS sire, group and AI sire equations are represented in matrix notation as:

$$\begin{bmatrix} \underset{\sim}{A} & \underset{\sim}{B} \\ \underset{\sim}{B}' & \underset{\sim}{D} \end{bmatrix} \begin{bmatrix} \hat{\underset{\sim}{g}} \\ \hat{\underset{\sim}{s}} \end{bmatrix} = \begin{bmatrix} \underset{\sim}{r}_1 \\ \underset{\sim}{r}_2 \end{bmatrix} \quad [2]$$

where $\hat{\underset{\sim}{g}}$ is a vector containing group solutions and the Lagrange multiplier and $\hat{\underset{\sim}{s}}$ is the vector of sire solutions.

### Variance of Prediction Errors

Prediction error variances can be obtained from the inverse of the coefficient matrix of the equations after absorption [2] (3).

$$\text{Var}\begin{bmatrix}\hat{\underset{\sim}{g}}\\ \hat{\underset{\sim}{s}}-\underset{\sim}{s}\end{bmatrix} = \begin{bmatrix}\underset{\sim}{A} & \underset{\sim}{B}\\ \underset{\sim}{B}' & \underset{\sim}{D}\end{bmatrix}^{-1}\sigma_e^2$$

$$= \begin{bmatrix}\underset{\sim}{C}_{11} & \underset{\sim}{C}_{12}\\ \underset{\sim}{C}'_{12} & \underset{\sim}{C}_{22}\end{bmatrix}\sigma_e^2 \qquad [3]$$

The inverse in most practical situations is difficult, if not impossible, to calculate by direct inversion techniques because of limitations on core storage. Therefore, an iterative procedure was used as described in (8, 9).

## RESULTS AND DISCUSSION

Two evaluations of sires were made; the first allowed up to 10 lactation records per cow, and the second allowed only first lactation records. Table 1 shows the counts including number of records, daughters, sires, and herds. The herd count is actually herds--not herd-year-seasons. The average of 2.5 lactations per cow is about one lactation less than the average life in the herd. This is probably due in part to the requirement that the first lactation record must be recorded in the same herd for records to be considered. Also, some cows did not have an opportunity to complete all their lactations before the herd went off test or before the data were cut off for analysis.

Inverse elements of the coefficient matrix were obtained by iteration after cow, sire-by-herd, NS sire, and herd-year-season equations were absorbed. Inverse elements give the error variances and covariances of prediction of the sire and group effect when multiplied by an appropriate $\sigma_e^2$. Sire evaluation is $(\hat{g}_i + \hat{s}_{ij})$.

From equation [3]:

$$V(\hat{g}_i) = (i, i\text{th element of } C_{11})\sigma_e^2 \qquad [4]$$

$$V(\hat{s}_{ij} - s_{ij}) = (ij, ij\text{th element of } C_{22})\sigma_e^2 \qquad [5]$$

$$\text{cov}(\hat{g}_i, \hat{s}_{ij} - s_{ij}) = (i, ij\text{th element of } C_{12})\sigma_e^2 \qquad [6]$$

$$V(\hat{g}_i + \hat{s}_{ij} - s_{ij}) = V(\hat{g}_i) + V(\hat{s}_{ij} - s_{ij}) + 2\text{cov}(\hat{g}_i, \hat{s}_{ij} - s_{ij}) \qquad [7]$$

The summary of inverse elements for Guernseys in Table 2 shows the relative contribution of variances of group and sire solutions and their covariances to error variances of evaluation. These elements are summarized by giving average values for AI sires that fall into various categories based on the number of daughters. The major contribution to the variance is the variance of the $(\hat{s}\text{-}s)$. The variance of $\hat{g}$ is less than one-fifth the variance for prediction errors with 20 to 50 daughters. The covariance is negative, and twice the covariance nearly cancels the contribution of the group variance. The pattern was similar for Ayrshires, Jerseys, and Brown Swiss (not shown).

TABLE 1. Numbers of records and effects from Northeast data.

| | Breed | | | |
|---|---|---|---|---|
| | Aryshire | Guernsey | Jersey | Brown Swiss |
| No. records | 63,271 | 111,150 | 111,890 | 23,309 |
| No. daughters | 23,831 | 44,692 | 43,393 | 9,181 |
| No. AI daughters | 14,642 | 23,699 | 23,446 | 5,571 |
| No. herds | 941 | 1,600 | 1,604 | 621 |
| No. sires by herds | 8,229 | 16,214 | 15,475 | 3,866 |
| No. AI sires | 122 | 376 | 315 | 113 |
| No. AI groups | 7 | 12 | 10 | 5 |
| No. NS groups | 8 | 8 | 8 | 8 |
| Herds/AI sire | 42 | 28 | 32 | 23 |
| AI and NS sires/herd | 8.7 | 10.1 | 9.6 | 6.2 |
| AI sires/herd | 5.4 | 6.5 | 6.3 | 4.2 |
| Records/daughter | 2.7 | 2.5 | 2.6 | 2.5 |

TABLE 2. Average variances and covariances of prediction error for AI group and sire solutions[a] categorized by number of daughters per sire for only first lactation records vs. all lactation records ($\sigma_e$ = 1)--Guernsey.

| Number of daughters | First lactations | | | | All lactations | | | |
|---|---|---|---|---|---|---|---|---|
| | $V(\hat{g})$ | $cov(\hat{g},\hat{s}-s)$ | $V(\hat{s}-s)$ | $V(\hat{g}+\hat{s}-s)$ | $V(\hat{g})$ | $cov(\hat{g},\hat{s}-s)$ | $V(\hat{s}-s)$ | $V(\hat{g}+\hat{s}-s)$ |
| 1 to 10 | .0136 | −.0015 | .1048 | .1154 | .0101 | −.0019 | .0974 | .1037 |
| 11 to 20 | .0127 | −.0039 | .0794 | .0843 | .0095 | −.0042 | .0665 | .0676 |
| 21 to 30 | .0095 | −.0039 | .0653 | .0670 | .0072 | −.0040 | .0504 | .0496 |
| 31 to 50 | .0091 | −.0045 | .0527 | .0528 | .0071 | −.0046 | .0398 | .0377 |
| 51 to 75 | .0098 | −.0056 | .0448 | .0434 | .0075 | −.0053 | .0328 | .0297 |
| 76 to 100 | .0107 | −.0071 | .0376 | .0341 | .0083 | −.0064 | .0278 | .0233 |
| 101 to 150 | .0103 | −.0072 | .0318 | .0277 | .0076 | −.0062 | .0225 | .0177 |
| 151 to 200 | .0064 | −.0047 | .0238 | .0207 | .0052 | −.0042 | .0173 | .0141 |
| 201 to 300 | .0097 | −.0075 | .0230 | .0177 | .0079 | −.0067 | .0176 | .0121 |
| 301+ | .0077 | −.0067 | .0152 | .0095 | .0062 | −.0057 | .0113 | .0061 |

[a]See equations [3], [4], [5], [6], and [7] in text for definitions of column headings.

**Prediction Errors of Group Solutions**

Standard errors of solutions for group effects are square roots of the variances described in [4]. The value of $\sigma_e$ was not obtained directly from this study. Relative values of standard errors are the primary interest in this study and are invariant to the value of $\sigma_e$ since $\sigma_e$ provides only a scaling factor. To put the standard errors on a scale that is reasonably appropriate for the breeds represented, $\sigma_e$ = 680 kg was chosen.

Standard errors of group solutions are in Table 3 through 6. Groups were set up to have a minimum of 10 sires per group. The AI groups were based on the AI stud and the year the sire entered service. Years and/or studs were

TABLE 3. Standard errors of group solutions for Ayrshires for only first lactation records vs. all lactation records ($\sigma_e$ = 680 kg).

| | No. sires | No. daughters | No. records | Standard errors | |
|---|---|---|---|---|---|
| | | | | First lactation | All lactations |
| | | | | (kg) | |
| AI group | | | | | |
| 1 | 21 | 1119 | 3606 | 83 | 72 |
| 2 | 18 | 7225 | 20873 | 70 | 65 |
| 3 | 19 | 1677 | 4521 | 70 | 64 |
| 4 | 20 | 1833 | 3189 | 70 | 64 |
| 5 | 11 | 449 | 1206 | 105 | 96 |
| 6 | 11 | 1153 | 2961 | 99 | 92 |
| 7 | 22 | 998 | 2563 | 84 | 73 |
| NS group | | | | | |
| 1 | | 1072 | 3197 | 62 | 45 |
| 2 | | 1119 | 3085 | 55 | 42 |
| 3 | | 1664 | 4712 | 48 | 37 |
| 4 | | 1519 | 4089 | 50 | 39 |
| 5 | | 1122 | 2990 | 54 | 42 |
| 6 | | 926 | 2383 | 59 | 46 |
| 7 | | 788 | 1720 | 62 | 49 |
| 8 | | 898 | 1253 | 58 | 47 |

TABLE 4. Standard errors of group solutions for Guernseys for only first lactation records vs. all lactation records ($\sigma_e$ = 680 kg).

| | No. sires | No. daughters | No. records | Standard errors | |
|---|---|---|---|---|---|
| | | | | First lactation | All lactations |
| | | | | (kg) | |
| AI group | | | | | |
| 1 | 43 | 4570 | 13276 | 59 | 52 |
| 2 | 44 | 4799 | 12979 | 50 | 45 |
| 3 | 47 | 3753 | 9296 | 48 | 43 |
| 4 | 38 | 1616 | 3276 | 57 | 51 |
| 5 | 19 | 394 | 1100 | 115 | 99 |
| 6 | 32 | 1080 | 2517 | 76 | 68 |
| 7 | 17 | 1531 | 3380 | 84 | 77 |
| 8 | 34 | 727 | 2142 | 86 | 68 |
| 9 | 29 | 1629 | 3994 | 72 | 62 |
| 10 | 24 | 591 | 1630 | 94 | 79 |
| 11 | 18 | 312 | 817 | 102 | 87 |
| 12 | 31 | 2513 | 5470 | 63 | 57 |
| NS group | | | | | |
| 1 | | 2126 | 5589 | 44 | 34 |
| 2 | | 2969 | 7816 | 39 | 31 |
| 3 | | 3584 | 9302 | 36 | 28 |
| 4 | | 3732 | 9278 | 34 | 28 |
| 5 | | 3050 | 7443 | 37 | 30 |
| 6 | | 2348 | 5440 | 40 | 32 |
| 7 | | 1629 | 3243 | 46 | 37 |
| 8 | | 1439 | 2068 | 49 | 39 |

combined to give adequate numbers of sires in each group. The AI sires were grouped essentially the same as sires in the Northeast AI Sire Comparison (NEAISC). The NS groups were based on year of registration. Year of registration was determined from a table that had the first registration number issued for each year for each breed. The intent was to group NS sires as nearly as possible on year of birth without actually having the birth date available. Each NS group covered 2 yr.

Standard errors of the sire evaluations will be the main basis for comparing procedures for first versus all lactation records; however, group standard errors decrease from only first to all lactation records. The addition of later lactation records reduced standard errors of the AI group solutions by 7 to 14% and of the NS group solutions by 20 to 24%. In addition, the effect number of sires, number of daughters, and number of records have on the standard errors can aid in forming sire groups.

**Prediction Errors of Sire Evaluations**

The main criterion for comparison of accuracy of the two alternative procedures is variances of prediction error of sire evaluations. The comparison of first only vs. all lactation records involves a subset of the data vs. all the data while the model remains the same. In this case variances can be derived directly from inverses of corresponding coefficient matrices. If the model changes from one procedure to the other, then the inverses might not yield the correct variances of prediction errors, and more effort would be required for a fair comparison of the variances under alternative models (4).

The model may appear different when only first lactations are considered because the cow effect, $c_{fijkm}$, no longer is needed. Fortunately, there are two equivalent ways of expressing the model. One is described with equation [1]. With only first lactation records, cow and error effects are confounded entirely but may be left separate and processed as though all lactations

TABLE 5. Standard errors of group solutions for Jerseys for only first lactation records versus all lactation records ($\sigma_e$ = 680 kg).

| | No. sires | No. daughters | No. records | Standard errors: First lactation | Standard errors: All lactations |
|---|---|---|---|---|---|
| | | | | (kg) | |
| AI group | | | | | |
| 1 | 37 | 2226 | 7150 | 62 | 55 |
| 2 | 43 | 4466 | 12754 | 47 | 43 |
| 3 | 39 | 6453 | 15622 | 48 | 44 |
| 4 | 33 | 1453 | 2814 | 59 | 54 |
| 5 | 23 | 1197 | 3194 | 97 | 87 |
| 6 | 13 | 1647 | 4746 | 104 | 96 |
| 7 | 31 | 929 | 2591 | 76 | 64 |
| 8 | 33 | 1216 | 2681 | 65 | 58 |
| 9 | 28 | 665 | 1764 | 89 | 74 |
| 10 | 35 | 2845 | 6499 | 57 | 52 |
| NS group | | | | | |
| 1 | | 1780 | 5032 | 47 | 37 |
| 2 | | 2157 | 6359 | 44 | 34 |
| 3 | | 2529 | 7089 | 40 | 32 |
| 4 | | 3292 | 8595 | 35 | 28 |
| 5 | | 2904 | 7912 | 38 | 30 |
| 6 | | 2644 | 6572 | 39 | 31 |
| 7 | | 2323 | 5343 | 42 | 33 |
| 8 | | 2154 | 3444 | 41 | 33 |

TABLE 6. Standard errors of group solutions for Brown Swiss for only first lactation records vs. all lactation records ($\sigma_e$ = 680 kg).

| | No. sires | No. daughters | No. records | Standard errors: First lactation | Standard errors: All lactations |
|---|---|---|---|---|---|
| | | | | (kg) | |
| AI group | | | | | |
| 1 | 22 | 1527 | 4890 | 81 | 67 |
| 2 | 25 | 2059 | 4661 | 71 | 62 |
| 3 | 16 | 470 | 1256 | 110 | 97 |
| 4 | 25 | 610 | 1558 | 83 | 72 |
| 5 | 25 | 846 | 1957 | 79 | 68 |
| NS group | | | | | |
| 1 | | 393 | 1121 | 102 | 75 |
| 2 | | 456 | 1310 | 93 | 68 |
| 3 | | 601 | 1581 | 84 | 63 |
| 4 | | 416 | 1086 | 89 | 68 |
| 5 | | 563 | 1391 | 81 | 63 |
| 6 | | 446 | 996 | 91 | 71 |
| 7 | | 410 | 786 | 96 | 74 |
| 8 | | 299 | 439 | 103 | 82 |

were present. The alternative is to define a new error term

$$\epsilon_{fijklm} = c_{fijkm} + e_{fijklm}$$

Then the model is rewritten as

$$y_{fijklm} = g_i + s_{fij} + h_{kl} + sh_{fijk} + \epsilon_{fijklm}$$

where the variance ratios are $\sigma^2_\epsilon/\sigma^2_s$ = 13.33 and $\sigma^2_\epsilon/\sigma^2_{sh}$ = 5.71. Identical sire predictions and variances of prediction error would result from either of these equivalent models for first lactation records.

Standard errors for sire evaluations were calculated as indicated in equation [7] with $\sigma_e$ = 680 kg. Average standard errors of prediction for sires falling in various categories based on number of daughters per sire are in Table 7. The average number of daughters for the sires that fell in each category is shown. Categories with 21 to 30 daughters and 31 to 59 daughters are ranges that include young sires eligible to be selected for return to service. Accuracy of evaluation of these sires has considerable influence on the theoretical rate of genetic trend. Standard errors are similar for all breeds. Brown Swiss is a slight exception with larger standard errors and greater reduction in standard errors from including the later records.

Sires were categorized further by number of records per daughter as illustrated for Guernseys in Table 8. In this table standard errors are reported with $\sigma_e$ = 1. Standard errors do not differ greatly as the number of records per daughter increases but appear to be more dependent on the number of daughters than the number of records per daughter. This was the pattern also for Ayrshires, Jerseys, and Brown Swiss (not shown) indicating the main reduction in error variance from first to all lactation records must be due to additional sire comparisons that are gained rather than the additional records per daughter. Additional sire comparisons arise because daughters of a sire are compared with daughters of other sires in herd-year-seasons where there were no comparisons for only first lactation records.

TABLE 7. Average standard errors of AI sire evaluations categorized by number of daughters per sire for only first lactation records vs. all lactations records ($\sigma_e$ = 680 kg).

| No. daughters range | Ayrshire | | | Guernsey | | | Jersey | | | Brown Swiss | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Ave. no. daughters | First lactation | All lactations | Ave. no. daughters | First lactation | All lactations | Ave. no. daughters | First lactation | All lactations | Ave. no. daughters | First lactation | All lactations |
| | | (kg) | | | (kg) | | | (kg) | | | (kg) | |
| 1 to 10 | 3 | 236 | 225 | 4 | 231 | 219 | 3 | 234 | 224 | 4 | 236 | 224 |
| 11 to 20 | 15 | 199 | 175 | 15 | 197 | 176 | 15 | 194 | 173 | 14 | 208 | 184 |
| 21 to 30 | 26 | 177 | 154 | 25 | 176 | 151 | 25 | 178 | 154 | 24 | 195 | 166 |
| 31 to 50 | 41 | 159 | 133 | 40 | 156 | 132 | 40 | 157 | 133 | 41 | 167 | 139 |
| 51 to 75 | 64 | 132 | 109 | 63 | 142 | 116 | 61 | 138 | 112 | 61 | 154 | 124 |
| 76 to 100 | 88 | 124 | 101 | 87 | 125 | 103 | 88 | 125 | 102 | 85 | 133 | 106 |
| 101 to 150 | 124 | 109 | 87 | 123 | 113 | 90 | 121 | 113 | 92 | 116 | 124 | 97 |
| 151 to 200 | 171 | 98 | 77 | 173 | 98 | 80 | 175 | 98 | 79 | 171 | 110 | 83 |
| 201 to 300 | 245 | 92 | 74 | 249 | 90 | 74 | 249 | 87 | 69 | 232 | 98 | 74 |
| 300+ | 1077 | 64 | 50 | 544 | 66 | 53 | 589 | 68 | 53 | 417 | 89 | 70 |

### Comparison of First Only versus All Lactation Records

At least two interpretations can be applied to differences in prediction errors between the

TABLE 8. Average standard errors of sire evaluations categorized by number of daughters per sire and by number of records per daughter for only first lactation records vs. all lactation records ($\sigma_e$ = 1)--Guernsey.

| | 1 Record/daughter | | | 1 to 2 Records/daughter | | | 2+ Records/daughter | | |
|---|---|---|---|---|---|---|---|---|---|
| No. daughters range | No. sires | First lactation | All lactations | No. sires | First lactation | All lactations | No. sires | First lactation | All lactations |
| 1 to 10 | 48 | .339 | .328 | 32 | .339 | .321 | 55 | .339 | .315 |
| 11 to 20 | 14 | .293 | .274 | 29 | .287 | .255 | 14 | .293 | .254 |
| 21 to 30 | 2 | .264 | .231 | 20 | .257 | .222 | 6 | .258 | .216 |
| 31 to 50 | 5 | .222 | .198 | 30 | .230 | .194 | 5 | .233 | .189 |
| 51 to 75 | 4 | .205 | .173 | 29 | .207 | .171 | 8 | .213 | .173 |
| 76 to 100 | 2 | .176 | .153 | 14 | .184 | .152 | 4 | .188 | .149 |
| 101 to 150 | 3 | .165 | .139 | 11 | .168 | .134 | 3 | .160 | .123 |
| 151 to 200 | 2 | .138 | .114 | 10 | .145 | .118 | 0 | .... | .... |
| 201 to 300 | 2 | .146 | .126 | 4 | .125 | .099 | 0 | .... | .... |
| 300+ | 2 | .109 | .090 | 15 | .094 | .075 | 1 | .112 | .086 |

two procedures. One is based on the change in numbers of daughters required for evaluations of similar accuracy. The other is concerned with effect on genetic trend.

An evaluation with 15 daughters per sire with all lactations gives nearly the same standard error as the evaluation with first lactations only of 25 daughters (Table 7). Similarly, an all-lactation evaluation with 25 daughters has about the same standard error as a first-lactation evaluation with 40 daughters, so that perhaps one-half to two-thirds more young sires might be tested on the same number of cows with comparable variances of prediction error if all lactations were used instead of only first lactations. This benefit would accrue with no increase in generation interval since use of later lactation records increased accuracy of evaluation of young sires with only one record per daughter nearly as much as for old sires with more than two records per daughter.

#### Expected Genetic Improvement

Expected genetic improvement may be a criterion to compare alternative strategies of evaluation. Comparison of expected genetic trends for BLUP evaluations requires simplifying assumptions. Sire evaluations would need to be normally and independently distributed with a single mean and the same variance of prediction error. These assumptions are probably reasonable for sires in contention for return to AI service following their initial proof. Any discrepancies from these assumptions probably affect both evaluation procedures in a similar manner, leaving the comparison of expected genetic trends valid.

The conventional equation for calculating expected genetic improvement per generation from truncation selection is given for selection index procedures as

$$\Delta G = r_{TI}\sigma_G D$$

where $r_{TI}$ is the correlation of the true genetic value with the selection index; D is the selection differential (height of ordinate of normal density function at truncation point divided by fraction selected); and $\sigma_G$ is the additive genetic standard deviation. Then, under the assumptions, improvement in daughter production may be expressed as

$$\Delta s = r_{\hat{s}s}\sigma_s D$$

The correlation, $r_{\hat{s}s}$, is not readily available so the equation may be rewritten as:

$$\Delta s = \text{cov}(\hat{s},s)\sigma_s D/\sigma_{\hat{s}}\sigma_s = \text{cov}(\hat{s},s)D/\sigma_{\hat{s}}$$

Then the BLUP properties $V(\hat{s}) = \text{cov}(\hat{s},s)$ and $\text{cov}(\hat{s},s) = V(s) - V(\hat{s}\text{-}s)$ (1) imply

$$\Delta s = [\text{cov}(\hat{s},s)]^{.5} D = [V(s) - V(\hat{s}\text{-}s)]^{.5} D \quad [8]$$

A choice must be made to use either prediction error variances for sire effects $V(\hat{s}\text{-}s)$ or error variances of evaluations $V(\hat{g}+\hat{s}\text{-}s)$. The former appears to more closely fit the assumptions for equation [8]. The selection of interest is for

young sires to return to service. With the stud-year grouping used in this study and in the NEAISC, selection would be mostly within groups, and the sire variance V($\hat{s}$-s) would be appropriate. However, some decisions in selection are between sires in different groups. Error variance of these selections is reflected in V($\hat{g}$+$\hat{s}$-s). Fortunately, the two values were nearly the same. Thus, for convenience error variances of Guernsey evaluations in Table 7 were substituted for V($\hat{s}$-s) in [8] to compare expected genetic progress for first and all lactations as in Table 9. The value of V(s) as determined from the ratio $\sigma_e^2/\sigma_s^2 = 8.33$ and $\sigma_e = 680$ is $236^2$. The estimate of relative progress for Guernsey sires with 40 daughters is 196/177 = 1.107; i.e., there is 10.7% more progress from additional records and 15.2% greater progress with 25 daughters.

Another important selection is the choice of sires and maternal grandsires for special matings to obtain young replacement sires. This selection is from all sires with semen available, including young, old, and dead sires, sires with few daughters, and sires with thousands of daughters. That is, selection is across genetic groups and involves unequal information on sires. This poses a problem for comparison of expected genetic progress from the two evaluation procedures. However, this selection normally will involve sires with as many or more records than the young sires selected for return to service. Table 9 shows that as the number of daughters per sire increases, the value of including the later lactation records decreases. Thus, the later lactation records are of less value in special matings. Results in Table 9 indicate that maximum increase in expected genetic progress is probably between 3 and 10%.

TABLE 9. Relative genetic progress from only first lactation records vs. all lactation records ($\sigma_e$ = 680 kg; standardized selection intensity factor, D = 1)–Guernsey.

| Ave. no. daughters | First lactation | All lactations | Ratio all/first |
|---|---|---|---|
| | (kg) | | |
| 4 | 48 D | 88 D | 1.840 |
| 15 | 129 D | 156 D | 1.211 |
| 25 | 157 D | 181 D | 1.153 |
| 40 | 177 D | 196 D | 1.105 |
| 63 | 188 D | 205 D | 1.087 |
| 87 | 200 D | 212 D | 1.059 |
| 123 | 206 D | 217 D | 1.053 |
| 173 | 214 D | 222 D | 1.034 |
| 249 | 218 D | 224 D | 1.027 |
| 544 | 226 D | 230 D | 1.016 |

The prediction of additional progress in Table 9 is based on the assumption that there is no bias introduced by later lactation records. Major sources of bias would be incorrect age factors, preferential treatment, and selection bias due to use of incorrect variance components. The most serious effect of these biases is on the comparison of young sires with older sires. While most selection tends to be within an age group (e.g., selection of young sires to return to service), there is some competition between age groups. Errors in age factors also affect selection within an age group. Daughters of some young sires may be compared with mostly first lactation records while daughters of other young sires may be compared with mostly later lactation records.

It is difficult, if not impossible, to determine the amount of bias in analyses of field data. However, some indication is possible in the present situation where the first lactation procedure is presumed unbiased. When later lactation records are included, changes from first lactation evaluations in various categories of sires is interpreted as an indication of bias. Keown et al. (5) studied trends in the sire evaluations. Including later lactation records caused young sires that entered service in 1967 to be biased upward 226 kg in Guernseys and 192 kg in Jerseys as compared to sires that entered service in 1950. Further, sires with more than two records per daughter were biased downward by 103 kg in Guernseys and 82 kg in Jerseys relative to sires with only one lactation per cow when additional records were used. In this study, the records had been age-season adjusted with an older set of age factors. Much of the bias could be attributed to these age factors and would be reduced greatly by improved age-season factors, e.g. (7). However, these results serve to indicate that biases may exceed the standard errors for many sires. Even with considerable care in estimating age-season factors and variance components, errors will occur and give rise to biases in sire evaluations when later lactation records are used.

On the whole, there appears to be value in using later lactation records if computing costs do not become prohibitive. Experience in this study indicates that computing costs are likely to be between 3 and 10 times the cost of using only first lactation records. It would appear desirable to run sire evaluations periodically, once every year or two, using only first lactation records to evaluate possible biases that may arise. Both first and all lactation evaluations could be published, making their strengths and weaknesses known. Even if the evaluations by first records only are not published, they would be helpful to AI studs for selective matings.

## ACKNOWLEDGMENTS

This research was supported in part by grants from Eastern Artificial Insemination Cooperative, Inc., Ithaca, NY 14850 and the USDA-ARS Animal Improvement Laboratory, Beltsville, MD 20705.

## REFERENCES

1 Henderson, C. R. 1973. Sire evaluation and genetic trends. Page 10 *in* Proc. Anim. Breeding and Genetics Symp. in Honor of Dr. Jay L. Lush, Amer. Soc. Anim. Sci., Champaign, IL.

2 Henderson, C. R. 1974. General flexibility of linear model techniques for sire evaluation. J. Dairy Sci. 57:963.

3 Henderson, C. R. 1975. Best linear unbiased estimation and prediction under a selection model. Biometrics 31:68.

4 Henderson, C. R. 1975. Comparison of alternative sire evaluation methods. J. Anim. Sci. 41:760.

5 Keown, J. R., G. R. Ufford, C. R. Henderson, and F. N. Dickinson. 1974. Sire evaluation for first lactations and all lactations from best linear unbiased prediction methods. J. Dairy Sci. 57:615 (Abstr.)

6 Norman, H. D. 1973. Personal communication. USDA-ARS, Beltsville, MD.

7 Norman, H. D., P. D. Miller, B. T. McDaniel, F. N. Dickinson, and C. R. Henderson. 1974. USDA-DHIA factors for standardizing 305-day lactation records for age and month of calving. Agr. Res. Serv., USDA ARS-NE-40.

8 Ufford, G. R., C. R. Henderson, and L. D. Van Vleck. 1979. An approximate procedure for determining prediction error variances of sire evaluation. J. Dairy Sci. (In press).

9 Ufford, G. R., C. R. Henderson, and L. D. Van Vleck. 1978. Derivation of computing algorithms for sire evaluation using all lactation records and natural service sires. Anim. Sci. Mimeo. Series No. 39.